PLPSTA31 – Statistiques 2
Tests d'hypothèses statistiques

Choix d'un test d'hypothèses statistiques
3 exercices

Mode d'emploi des questionnaires à choix multiples

Pour démarrer un exercice, il faut cliquer sur Début.
Pour afficher le score obtenu au questionnaire et les réponses correctes, il faut cliquer sur Fin, puis sur le bouton Réponse :

- le signe ✔ indique que la réponse donnée est correcte ;
- le signe ✘ indique que la réponse donnée est incorrecte ; les réponses correctes manquantes sont marquées par ●.

Pour certaines questions, les réponses sont détaillées : cliquer sur un des symboles verts ✔ ou ● pour accéder aux compléments de réponse.
Attention : il peut y avoir plusieurs propositions correctes dans une liste. Il faut toutes les cocher pour que la réponse à la question soit considérée comme exacte.

Exercice 1

Le niveau de symptômes de stress post-traumatique chez des enfants victimes d'un accident domestique ou de la circulation est mesuré par un score (score PTSD) ; ce score augmente avec la sévérité des symptômes.

On cherche à évaluer l'efficacité d'une thérapie chez ces enfants. Sur un échantillon d'enfants victimes ayant bénéficié de cette thérapie, on relève : le sexe (S), le type d'accident survenu (A), le score avant thérapie (T_1), le score après thérapie (T_2) et la différence de scores $T_1 - T_2$.

Pour chacune des questions suivantes, il faut identifier la ou les variable(s) étudiée(s) ainsi que le test d'hypothèses à réaliser.

Début

1. Y a-t-il une diminution significative du niveau de symptômes en fin de thérapie tous types d'accidents confondus ?

(a) Variable(s) étudiée(s) : ☐ S ☐ A ☐ T_1 ☐ T_2 ☐ $T_1 - T_2$

(b) Test à utiliser :

☐ test sur une proportion ☐ test du khi-deux d'ajustement ☐ test du khi-deux d'homogénéité
☐ test du khi-deux d'indépendance ☐ test sur une moyenne ☐ test de comparaison de 2 moyennes

2. Chez les enfants victimes d'un accident, existe-t-il un lien entre le sexe et le type d'accident survenu ?

(a) Variable(s) étudiée(s) : ☐ S ☐ A ☐ T_1 ☐ T_2 ☐ $T_1 - T_2$

(b) Test à utiliser :

☐ test sur une proportion ☐ test du khi-deux d'ajustement ☐ test du khi-deux d'homogénéité
☐ test du khi-deux d'indépendance ☐ test sur une moyenne ☐ test de comparaison de 2 moyennes

3. Le niveau de symptômes avant thérapie est-il plus élevé chez les enfants victimes d' un accident de la circulation que chez les victimes d'un accident domestique ?

(a) Variable(s) étudiée(s) : ☐ S ☐ A ☐ T_1 ☐ T_2 ☐ $T_1 - T_2$

(b) Test à utiliser :

☐ test sur une proportion ☐ test du khi-deux d'ajustement ☐ test du khi-deux d'homogénéité
☐ test du khi-deux d'indépendance ☐ test sur une moyenne ☐ test de comparaison de 2 moyennes

Fin Score : Réponse Cliquer sur ✔ ou ● pour accéder aux compléments de réponse.

Exercice 2

Trois échantillons de patients en consultation dans une clinique psychiatrique ont été constitués :

un échantillon de 101 patients atteints de troubles paniques (TP), un échantillon de 59 patients atteints d'autres troubles d'anxiété (TA) et un échantillon de 20 patients subissant des épisodes dépressifs majeurs récurrents (EDM).

Sur chacun des 180 patients on a relevé : l'âge auquel les premiers troubles sont apparus ; les antécédents de violence subie pendant l'enfance, variable à 2 modalités *oui*, *non*.

Pour chacune des questions suivantes, indiquer la ou les population(s) étudiée(s) ainsi que le test d'hypothèses à réaliser.

Début

1. Chez les patients atteints de troubles paniques, les premiers troubles apparaissent-ils plus tôt chez ceux ayant subi des violences précoces ?

(a) Population(s) : ☐ « TP » ☐ « TA » ☐ « EDM » ☐ « TP » avec violences subies ☐ « TP » sans violences subies

(b) Test à effectuer :

☐ test sur une proportion ☐ test du khi-deux d'ajustement ☐ test du khi-deux d'homogénéité
☐ test du khi-deux d'indépendance ☐ test sur une moyenne ☐ test de comparaison de 2 moyennes

2. On se demande si moins de 50% des patients atteints de troubles paniques ont subi des violences pendant leur enfance.

(a) Population(s) : ☐ « TP » ☐ « TA » ☐ « EDM » ☐ « TP » avec violences subies ☐ « TP » sans violences subies

(b) Test à effectuer :

☐ test sur une proportion ☐ test du khi-deux d'ajustement ☐ test du khi-deux d'homogénéité
☐ test du khi-deux d'indépendance ☐ test sur une moyenne ☐ test de comparaison de 2 moyennes

3. Existe-t-il une différence entre les trois types de patients concernant les antécédents de violences subies pendant l'enfance ?

(a) Population(s) : ☐ « TP » ☐ « TA » ☐ « EDM » ☐ « TP » avec violences subies ☐ « TP » sans violences subies

(b) Test à utiliser :

☐ test sur une proportion ☐ test du khi-deux d'ajustement ☐ test du khi-deux d'homogénéité
☐ test du khi-deux d'indépendance ☐ test sur une moyenne ☐ test de comparaison de 2 moyennes

Fin | Score : | Réponse   Cliquer sur ✔ ou ● pour accéder aux compléments de réponse.

Exercice 3

On connaît (*source INSEE*) la distribution des départs en vacances des ménages français en 1999 :

pas de vacances (N)	uniquement en été (E)	uniquement en hiver (H)	en été et en hiver (E& H)
38,5%	34%	4,5%	23%

En 2004, sur un échantillon de 1000 ménages français réunissant 3 échantillons indépendants de ménages appartenant à des tranches de revenu différentes, on a observé les effectifs :

N	E	H	E&H
354	322	53	271

Pour les 230 ménages français de cet échantillon appartenant à la tranche inférieure de revenu, on a observé les effectifs :

N	E	H	E&H
124	71	12	23

Pour les 180 ménages français de cet échantillon appartenant à la tranche supérieure de revenu, on a observé les effectifs :

N	E	H	E&H
29	50	10	91

Pour chacune des questions suivantes, il faut identifier le test d'hypothèses à réaliser.

Début

1. La distribution des départs en vacances de 2004 diffère-t-elle de celle de 1999 ?

- ☐ test sur une proportion
- ☐ test du khi-deux d'ajustement
- ☐ test du khi-deux d'homogénéité
- ☐ test du khi-deux d'indépendance
- ☐ test sur une moyenne
- ☐ test de comparaison de 2 moyennes

2. Le taux de non-départ en vacances en 2004 est-il inférieur à celui de 1999 ?

- ☐ test sur une proportion
- ☐ test du khi-deux d'ajustement
- ☐ test du khi-deux d'homogénéité
- ☐ test du khi-deux d'indépendance
- ☐ test sur une moyenne
- ☐ test de comparaison de 2 moyennes

3. En 2004, le taux de non-départ des ménages français de la tranche inférieure de revenu est-il supérieur à 50% ?

- ☐ test sur une proportion
- ☐ test du khi-deux d'ajustement
- ☐ test du khi-deux d'homogénéité
- ☐ test du khi-deux d'indépendance
- ☐ test sur une moyenne
- ☐ test de comparaison de 2 moyennes

4. En 2004, la distribution des départs en vacances de la tranche inférieure de revenus diffère-t-elle de celle de la tranche supérieure de revenus ?

- ☐ test sur une proportion
- ☐ test du khi-deux d'ajustement
- ☐ test du khi-deux d'homogénéité
- ☐ test du khi-deux d'indépendance
- ☐ test sur une moyenne
- ☐ test de comparaison de 2 moyennes

Fin | Score : | Réponse   Cliquer sur ✔ ou ● pour accéder aux compléments de réponse.

Retour

Les réponses aux questionnaires

Réponse :

On étudie ici la population des enfants victimes d'un accident domestique ou de la circulation ayant bénéficié de la thérapie. Pour évaluer l'évolution du niveau de symptômes des enfants entre le début et la fin de la thérapie, on utilise la variable quantitative *Différence des scores* $T_1 - T_2$.

Cette variable a une moyenne (différence de scores moyenne de la population) μ *inconnue*.

Retour au questionnaire.

Réponse :

Le test porte sur la moyenne *inconnue* μ de la variable $T_1 - T_2$.

L'hypothèse d'une diminution du niveau de symptômes en fin de thérapie se traduit globalement sur la population par des scores PTSD en baisse, c'est-à-dire par des différences de scores positives et donc par une différence de scores moyenne $\mu > 0$.

On doit donc tester $H_0 : \mu = 0$ contre $H_1 : \mu > 0$. Le test est unilatéral droit.

Selon la distribution de la variable $T_1 - T_2$ (normale ou inconnue),il y a deux procédures possibles : le test de Student ou le test basé sur l'approximation normale (si l'échantillon est assez grand). Retour au questionnaire.

Réponse : La population étudiée est la population des enfants victimes d'un accident domestique ou de la circulation. On s'intéresse au sexe et au type d'accident survenu.

On étudie ici les 2 variables *Sexe* (S) et *Type d'accident* (A) qui sont toutes les deux qualitatives à 2 modalités.

Retour au questionnaire.

Retour

Réponse :

Pour tester l'existence d'une liaison entre 2 variables qualitatives, on utilise le test du khi-deux d'indépendance.

On teste H_0 : les variables sont indépendantes contre H_1 : les 2 variables sont liées.

La procédure est basée sur la loi du khi-deux à un degré de liberté.
Pour appliquer le test, on doit au préalable vérifier que les conditions d'application du test sont satisfaites.

Retour au questionnaire.

Retour

Réponse :

On veut comparer le niveau de symptômes avant thérapie chez les victimes d'un accident domestique et chez les enfants victimes d'un accident de la circulation.

On étudie ici deux populations d'enfants :

$\mathscr{P}_1$: les enfants victimes d'un accident domestique et $\mathscr{P}_2$: les enfants victimes d'un accident de la population.

Le niveau de symptômes avant thérapie est mesuré par la variable quantitative T_1. Elle est mesurée sur les deux populations. En suivant les notations du cours, on peut la noter X_1 sur $\mathscr{P}_1$ et X_2 sur $\mathscr{P}_2$.

La moyenne μ_1 de X_1 représente le score moyen avant thérapie dans $\mathscr{P}_1$ et la moyenne μ_2 de X_2 le score moyen avant thérapie dans $\mathscr{P}_2$. Ces deux moyennes sont *inconnues*.

Retour au questionnaire.

Réponse :

On effectue ici un test de comparaison de deux moyennes à partir de deux échantillons indépendants. L'échantillon initial est divisé en deux échantillons indépendants : un échantillon d'enfants victimes d'un accident domestique et un échantillon d'enfants victimes d'un accident de la circulation.

L'hypothèse d'un niveau de symptômes plus élevé chez les enfants victimes d'un accident de la circulation se traduit par l'hypothèse unilatérale $H_1 : \mu_1 < \mu_2$. L'hypothèse nulle est $H_0 : \mu_1 = \mu_2$.

Comme pour le test sur une moyenne, il y a deux procédures possibles : test de Student ou test basé sur l'approximation normale. Chaque procédure a ses propres conditions d'application.

Retour au questionnaire.

Réponse :

On veut comparer l'âge d'apparition des premiers troubles chez les patients « TP »ayant subi des violences au cours de l'enfance (population $\mathscr{P}_1$) et chez les patients « TP »n'en ayant pas subi (population $\mathscr{P}_2$).

La variable quantitative *Age d'apparition des troubles* sera notée X_1 sur $\mathscr{P}_1$ et X_2 sur $\mathscr{P}_2$. Retour au questionnaire.

Réponse :

On compare les âges moyens d'apparition des troubles des deux populations.

La moyenne μ_1 de X_1 représente l'âge moyen d'apparition des troubles chez les patients TP ayant subi des violences pendant l'enfance.

La moyenne μ_2 de X_2 représente l'âge moyen d'apparition des troubles chez les patients TP n'ayant pas subi de violences pendant l'enfance.

On effectue ici un test de comparaison de deux moyennes à partir de deux échantillons indépendants.

La question posée est traduite par l'hypothèse unilatérale $H_1 : \mu_1 < \mu_2$. L'hypothèse nulle s'écrit $H_0 : \mu_1 = \mu_2$.

L'échantillon « TP » initial est divisé en deux échantillons indépendants : un échantillon de patients TP ayant subi des violences et un échantillon de patients TP n'ayant pas subi de violences.

Retour au questionnaire.

Retour

Réponse :

La population étudiée est celle des patients « TP ». Dans cette population, on s'intéresse aux patients qui ont subi des violences pendant leur enfance. Cette catégorie de patients est codée par la modalité *oui* de la variable qualitative *Antécédents de violence subie pendant l'enfance*. Retour au questionnaire.

Réponse :

Dans la population « TP », on s'interroge sur la proportion p *inconnue* de la catégorie de patients ayant subi des violences pendant leur enfance.

On fait un test de proportion pour comparer p à 50% (valeur théorique).

On teste $H_0 : p = 0,5$ contre $H_1 : p < 0,5$. Le test est unilatéral gauche.

Pour appliquer le test, on doit au préalable vérifier que les conditions d'application du test sont satisfaites.

Retour au questionnaire.

Retour

Réponse :

On compare ici les distributions *inconnues* de la variable *Antécédents de violence subie pendant l'enfance* dans les 3 populations de patients « TP », « TA » et « EDM ».

Pour faire le test, on dispose de 3 échantillons indépendants : un échantillon de patients « TP », un échantillon de patients « TA » et un échantillon de patients « EDM ».

Il s'agit d'un test du khi-deux d'homogénéité.

Les hypothèses à tester sont :

H_0 : la variable *Antécédents de violence...* a la même distribution dans les trois populations de patients (ou *le pourcentage de patients ayant subi des violences pendant leur enfance est identique dans les 3 populations*)

et

H_1 : la variable *Antécédents de violence...* n'a pas la même distribution dans les trois populations de patients.

Le test utilise la loi du khi-deux à 2 ddl. Les conditions d'application du test sont à vérifier.

Retour au questionnaire.

Retour

Réponse :

Préalable :

Toutes les questions portent sur des variables qualitatives. Les tests sur une ou 2 moyenne(s) n'ont pas de sens pour cet exercice.

Il s'agit de décider si la distribution *connue* des départs en vacances de 1999 peut s'appliquer aux ménages français de 2004. Cette distribution est la distribution théorique testée.

Le test du khi-deux d'ajustement permet de comparer la distribution *inconnue* des départs en vacances de 2004 à la distribution théorique établie en 1999.

La population étudiée est composée des ménages français de 2004.

La variable *départs en vacances* étudiée est qualitative à 4 modalités N, E, H, $E\&H$.

Les hypothèses à tester sont :

H_0 : la distribution des départs en vacances de 2004 est identique à la distribution des départs en vacances de 1999, soit :

N : 38,5% ; E : 34% ; H : 4,5% et E&H : 23%

H_1 : la distribution des départs en vacances de 2004 diffère de la distribution des départs en vacances de 1999

Le test utilise les effectifs n_i

N	E	H	E&H
354	322	53	271

observés sur l'échantillon de taille 1000 tiré au sort dans la population.

Le test est basé sur la loi du khi-deux à 3 ddl.

Retour au questionnaire.

Retour

Réponse :

La population est composée des ménages français de 2004.

La variable *qualitative* étudiée ici permet de diviser la population en deux catégories : la catégorie *non-départ* (codée N) et la catégorie *départ* regroupant tous ceux qui partent en vacances.

On s'interroge sur la proportion p *inconnue* de ménages français de 2004 n'ayant pas pris de vacances (proportion de la catégorie N).

On compare p à la proportion *connue* (38,5%) de la même catégorie en 1999.

Les hypothèses à tester sont : $H_0 : p = 0,385$ et $H_1 : p < 0,385$. Le test est unilatéral gauche.

Le test utilise la fréquence de non-départ observée sur l'échantillon de taille 1000 : $f_{\text{obs}} = \frac{354}{1000} = 0,354$.

Retour au questionnaire.

Retour

Réponse :

La population est composée des ménages français de 2004 appartenant à la tranche inférieure de revenu.
Ici, on note p la proportion *inconnue* de non-départ dans cette population. On compare p à la valeur théorique 50%.

Les hypothèses à tester sont : $H_0 : p = 0,50$ et $H_1 : p > 0,50$. Le test est unilatéral droit.

Sur l'échantillon de taille 230 tiré au sort dans la population, la fréquence de non-départ observée est $f_{\text{obs}} = \frac{124}{230} = 0,539$.

Retour au questionnaire.

Retour

Réponse :

On compare ici les distributions *inconnues* des départs en vacances de 2004 pour les deux tranches de revenu *inférieure* et *supérieure*.

Il s'agit d'un test du khi-deux d'homogénéité.

On étudie la variable *Départs en vacances* (à 4 modalités) dans deux populations :

- la population de ménages français de 2004 appartenant à la tranche inférieure de revenu ;
- la population de ménages français de 2004 appartenant à la tranche supérieure de revenu.

Les hypothèses à tester sont :

H_0 : la distribution des départs en vacances en 2004 est identique dans les deux tranches de revenu
et
H_1 : la distribution des départs en vacances en 2004 diffère selon la tranche de revenu

On dispose de deux échantillons indépendants : un échantillon de 230 ménages de 2004 appartenant à la tranche inférieure de revenu et un échantillon de 180 ménages de 2004 appartenant à la tranche supérieure.

Pour le test, on utilise le tableau des effectifs observés n_{ij} :

	Tranche de revenu	
Départs en vacances	inférieure	supérieure
N	124	29
E	71	50
H	12	10
E&H	23	91

Le test est basé sur la loi du khi-deux à 3 ddl.

Retour au questionnaire.

Retour